Cumulating samples

The remaining option on the Pre-treatment menu is **Cumulate samples**, which successively adds up the entries across variables, separately for each sample. It is only appropriate when all variables share a common measurement scale, and when the order in which they are listed is meaningful; it is thus not relevant standard species-by-samples data. It may be useful in analysing arrays in which variables are different body-size categories of a single species, or different particle sizes classes in Particle Size Analysis (PSA) etc, and entries are the frequencies or quantities of each size class in each sample (see CiMC, Chapter 8). Such data is typically analysed by univariate methods, fitting parametric particle-size distributions and comparing parameter estimates over samples. That can be problematic: histograms do not fit the models, summary statistics like mean and variance do not capture features such as bimodality, tests are incorrect because the data are not real frequencies, it is difficult to synthesise many such samples etc. This can be side-stepped by multivariate analysis, defining the similarity of pairs of size-class distributions. To take into account ordering of the sizes, when histograms are not smooth, it may sometimes be preferable to compare pairs of cumulative distributions (sample *distribution functions*) rather than the histograms (sample *density functions*).

Revision #1
Created 22 May 2024 01:23:26 by Arden
Updated 22 May 2024 01:25:32 by Arden